Discrete Applied Mathematics 7 (1984) 161-181
North-Holland



# ON THE AVERAGE SHAPE OF MONOTONICALLY LABELLED TREE STRUCTURES

Peter KIRSCHENHOFER

*Institut für Algebra und Diskrete Mathematik, Technische Universität Wien, Gusshausstr. 27–29, A-1040 Vienna, Austria*

Received 21 August 1981
Revised 17 December 1982

Let $\mathscr{B}_k$ denote one of the families of binary trees, $t$-ary trees ($t>2$) or ordered trees with nodes labelled monotonically by elements of $\{1<2<\cdots<k\}$. The average height of the $j$-th leaf of the trees of $\mathscr{B}_k$ with exactly $n$ nodes is shown to converge to a finite limit (depending on $k$ and $j$) for $n\to\infty$. The limit is determined explicitly for small values of $k$ and its asymptotic behaviour in $j$ and $k$ is investigated. Some recent results on the average shape of rooted tree structures appear as special cases.

## 1. Introduction

Consider a rooted tree structure (in the sense of Knuth [7]) the nodes of which have been labelled monotonically by elements of $\{1<2<\cdots<k\}$, which means that any sequence of labels starting from the root of the tree is weakly monotone. In a recent paper [9] Prodinger and Urbanek have considered the problem of finding asymptotic equivalents to the numbers of such tree structures with $n$ nodes in the case of some special families such as binary trees, $t$-ary trees, ordered trees etc. In the present work we want to give information on the average shape of some of these families by results of the following type:

*Let $\mathscr{B}_k$ be a given family of rooted tree structures labelled monotonically by elements of $\{1<2<\cdots<k\}$. The average height of the $j+1$-st leaf of the trees of $\mathscr{B}_k$ with exactly n nodes (where all such trees of $\mathscr{B}_k$ are regarded equally likely and leaves are enumerated from left to right) is shown to converge to a finite limit*

$$\alpha_k(j) \quad \textit{for } n\to\infty.$$

*(Speaking less rigorously we will say, $\alpha_k(j)$ describes the average height of the $j+1$-st leaf for 'large node number n'.)*

*Studying a well-suited generating function of the numbers $\alpha_k(j)$ allows us*

(a) *to compute $\alpha_k(j)$ for small values of $k$ explicitly,*

(b) *to establish asymptotic results of the type*

$$\alpha_k(j)\sim C_k j^{1/2} \qquad (j\to\infty)$$

*where the constants $C_k$ are expressed in terms of some 'characteristic quantities' of the families $\mathscr{B}_k$, which have been studied extensively in the paper* [9]. *This allows us to describe the asymptotic behaviour of $C_k$ as*

$$C_k \sim Ck^{-1/2} \qquad (k \to \infty)$$

*with a constant $C$ independent from $k$.*

Some of the material on the average shape of trees which has been worked out recently appears in special cases of results presented in these papers, e.g.

(a) A theorem of Ruskey [10] on binary trees, which in our terminology reads

$$\alpha_1(j) = 2 \cdot 4^{-j}(j+1)\binom{2j+2}{j+1} - 1$$
$$\sim 8\pi^{-1/2}j^{1/2} \qquad (j \to \infty).$$

(b) A result of Kemp in his investigation [5] on the average oscillations of a stack during postorder-traversing of a binary tree. The average heights of the so called 'MAX-turns' of the stack correspond to our numbers $\alpha_1(j)$ where $\mathscr{B}$ is the family of ordered trees and fulfill the asymptotic formula

$$\alpha_1(j) \sim 8(2\pi)^{-1/2}j^{1/2} \qquad (j \to \infty).$$

In Section 2 we will try to give a relatively detailed development of our method in the case of binary trees. The method proves to be transferable to $t$-ary trees for general $t \geq 2$ (where the results are less 'legible' than in the special case $t=2$, which was the main reason for treating that case at first) and ordered trees. This will be done in Sections 3 and 4, where the style of presentation will be shorter for the sake of brevity of the whole paper.

**Remark.** In the following considerations we will frequently use the abbreviation $\langle f, z^n \rangle$ for the coefficient of $z^n$ in the (formal) power series $f(z)$.

## 2. Binary trees

Let $\mathscr{B}_k$ denote the family of extended binary trees (in the sense of Knuth [7]) labelled monotonically by elements of $\{1<2<\cdots<k\}$ and $\tilde{\mathscr{B}}_k$ the corresponding family with labels taken from $\{2<\cdots<k+1\}$. In the suggestive terminology of Flajolet (compare [3], and the comments in [9]) the families $\mathscr{B}_k$ may be defined by the formal equations

$$\mathscr{B}_1 = \square + \begin{array}{c} (1) \\ \diagup \quad \diagdown \\ \mathscr{B}_1 \qquad \mathscr{B}_1 \end{array}$$

$$\begin{aligned}
\mathcal{B}_2 &= \square + \begin{array}{c} \textcircled{2} \\ \tilde{\mathcal{B}}_1 \quad \tilde{\mathcal{B}}_1 \end{array} + \begin{array}{c} \textcircled{1} \\ \mathcal{B}_2 \quad \mathcal{B}_2 \end{array} = \tilde{\mathcal{B}}_1 + \begin{array}{c} \textcircled{1} \\ \mathcal{B}_2 \quad \mathcal{B}_2 \end{array} \\
&\dots \\
\mathcal{B}_k &= \tilde{\mathcal{B}}_{k-1} + \begin{array}{c} \textcircled{1} \\ \mathcal{B}_k \quad \mathcal{B}_k \end{array} \\
&\dots
\end{aligned} \tag{2.1}$$

where $\square$ is the symbol for a leaf and $\bigcirc$ for an (internal) node of the binary tree.

Let $\mathcal{B}_{k,n}$ denote the family of trees in $\mathcal{B}_k$ with exactly $n$ (internal) nodes and $y_k = \sum_n \langle y_k, z^n \rangle z^n$ the generating function of the numbers $\langle y_k, z^n \rangle$ of trees in $\mathcal{B}_{k,n}$. (2.1) yields

$$y_k = y_{k-1} + z y_k^2, \quad k \ge 1; \qquad y_0 = 1 \tag{2.2}$$

(compare [9]). Let $t$ be a tree of $\mathcal{B}_k$ with at least $j+1$ leaves ($j \ge 0$). By $h_j(t)$ we denote the height of the $j+1$-st leaf, that means the number of elements of the chain connecting the root with the $j+1$-st leaf where the leaf itself is *not* counted.

If all trees of $\mathcal{B}_{k,n}$ are regarded equally likely, the average height of the $j+1$-st leaf is given by

$$\langle y_k, z^n \rangle^{-1} \sum_{t \in \mathcal{B}_{k,n}} h_j(t) \qquad (n \ge j). \tag{2.3}$$

So we introduce the generating functions

$$H_{k,j}(z) := \sum_{n \ge j} z^n \sum_{t \in \mathcal{B}_{k,n}} h_j(t). \tag{2.4}$$

Let $v_{k,j}^{[h]}(z)$ denote the generating function of the numbers $\langle v_{k,j}^{[h]}(z), z^n \rangle$ of trees $t$ of $\mathcal{B}_{k,n}$, $n \ge j$, with $h_j(t) \le h$ and

$$y_{k,j}(z) := \sum_{n \ge j} \langle y_k, z^n \rangle z^n \tag{2.5}$$

the generating function of the numbers of trees with at least $j+1$ leaves. Then, as an easy consequence of the definitions,

$$H_{k,j} = \sum_{h \ge 0} (y_{k,j} - v_{k,j}^{[h]}). \tag{2.6}$$

In order to derive a recurrence relation for the functions $H_{k,j}$ we first observe that by (2.2) and (2.5)

$$y_{k,j} = y_{k-1,j} + z(y_k^2)_{j-1}, \quad j \ge 1; \qquad y_{0,j} = \delta_{0,j} \tag{2.7}$$

with

$$\begin{aligned}
(y_k^2)_{j-1} &:= \sum_{n \ge j-1} z^n \langle y_k^2, z^n \rangle \\
&= \sum_{i \ge j} \langle y_k, z^i \rangle z^i \sum_{n \ge i} \langle y_k, z^{n-i} \rangle z^{n-i} \\
&\quad + \sum_{i=0}^{j-1} \langle y_k, z^i \rangle z^i \sum_{n \ge j-1} \langle y_k, z^{n-i} \rangle z^{n-i}.
\end{aligned}$$

So we get

$$y_{k,j} = y_{k-1,j} + zy_{k,j}y_k + z\sum_{i=0}^{j-1} \langle y_k, z^i\rangle z^i y_{k,j-1-i}. \tag{2.8}$$

A similar relation holds for the functions $v_{k,j}^{[h]}$:

$$v_{k,j}^{[h+1]} = v_{k-1,j}^{[h+1]} + zv_{k,j}^{[h]}y_k + z\sum_{i=0}^{j-1} \langle y_k, z^i\rangle z^i v_{k,j-1-i}^{[h]} \quad \text{for all } j,h \geq 0. \tag{2.9}$$

This can be seen by observing that a tree $t \in \mathscr{B}_k$ with $h_j(t) \leq h+1$ may be $\square$ or have a root labelled by elements of $\{2, \ldots, k\}$, which yields the first term of the sum, or it starts with a root labelled by 1. In the latter case the $j+1$-st leaf can be situated in the left subtree, which means that there is no restriction on the right subtree and yields the second term, or it lies in the right subtree, which means that the left subtree must have less than $j$ inner nodes and establishes the third term.

Subtracting relations (2.8) and (2.9) and summing up over all $h \geq 0$ gives

$$H_{k,j} - (y_{k,j} - v_{k,j}^{[0]}) = H_{k-1,j} - (y_{k-1,j} - v_{k-1,j}^{[0]}) + zH_{k,j}y_k + z\sum_{i=0}^{j-1} \langle y_k, z^i\rangle z^i H_{k,j-1-i}$$

and because of $v_{i,j}^{[0]} = \delta_{j,0}$ for all $i \geq 0$, we get the desired recurrence

$$H_{k,j}(1 - zy_k) = H_{k-1,j} + (y_{k,j} - y_{k-1,j}) + z\sum_{i=0}^{j-1} \langle y_k, z^i\rangle z^i H_{k,j-1-i} \tag{2.10}$$

with $H_{0,j} = 0$.

Let $q_k$ be the radius of convergence of the function $y_k$. In [9] it is shown that the numbers $q_k$ fulfill the recursion

$$q_{k+1} = q_k(1 - q_k) \quad \text{with } q_1 = \tfrac{1}{4}, \tag{2.11}$$

that $z = q_k$ is the only singularity on the circle of convergence of $y_k$ and $y_k$ behaves like

$$y_k(z) = y_k(q_k) - a_k(q_k - z)^{1/2} + \mathcal{O}(q_k - z) \qquad (z \to q_k^-). \tag{2.12}$$

Of course the functions $y_{k,j}$ have the same singularities as $y_k$ and behave like

$$y_{k,j}(z) = y_{k,j}(q_k) - a_k(q_k - z)^{1/2} + \mathcal{O}(q_k - z) \tag{2.13}$$

where $a_k$ is the same constant as in (2.12).

Consider now recursion (2.10) again: By (2.2)

$$(1 - zy_k)^{-1} = y_k / y_{k-1}$$

and because of $q_k < q_{k-1} < \cdots < q_1$, $H_{k,j}$ must have its singularity nearest to the origin at $z = q_k$, too, and, again by (2.10) and (2.13), behave like

$$H_{k,j}(z) = H_{k,j}(q_k) - a_{k,j}(q_k - z)^{1/2} + \mathcal{O}(q_k - z). \tag{2.14}$$

The asymptotic behaviour of the coefficients

$$\langle H_{k,j}, z^n\rangle = \sum_{t \in \mathscr{B}_{k,n}} h_j(t) \qquad (n \geq j)$$

of $H_{k,j}$ (compare (2.4)) follows by a theorem of Darboux (see e.g. [2, p. 277], [4, p. 211f.], [8] or the applications in [9]) to be

$$\langle H_{k,j}, z^n\rangle \sim \frac{a_{k,j}}{2}\left(\frac{q_k}{\pi}\right)^{1/2} q_k^{-n} n^{-3/2} \qquad (n \to \infty) \tag{2.15}$$

and because of

$$\langle y_k, z^n\rangle \sim \frac{a_k}{2}\left(\frac{q_k}{\pi}\right)^{1/2} q_k^{-n} n^{-3/2}$$

(which again follows by Darboux's theorem from relation (2.12), compare also [9]), we get

$$\alpha_k(j) := \lim_{n\to\infty} \frac{\langle H_{k,j}, z^n\rangle}{\langle y_k, z^n\rangle} = \frac{a_{k,j}}{a_k} \tag{2.16}$$

where $\alpha_k(j)$ denotes the desired average height of the $j+1$-st leaf for 'large $n$' (compare the comments in the introduction). Let

$$A_k(u) := \sum_{j\geq 0} \alpha_k(j) u^j \tag{2.17}$$

be the corresponding generating function in $u$ and

$$H_k(z,u) := \sum_{j \geq 0} H_{k,j}(z) u^j. \tag{2.18}$$

Relations (2.16) for all $j \geq 0$ may be rewritten now in shorter form as

$$a_k A_k(u) = \lim_{z \to q_k^-} \frac{H_k(q_k, u) - H_k(z,u)}{(q_k - z)^{1/2}}, \tag{2.19}$$

where the limit on the right side of the equation is to be carried out for each coefficient of the term, which is considered as a formal power series in $u$.

A recurrence relation on the series $H_k(z,u)$ is derived in the following way: By (2.10) we have

$$\begin{aligned} H_k(z,u)(1 - zy_k(z)) = H_{k-1}(z,u) &+ \sum_{j\geq 0} u^j (y_{k,j}(z) - y_{k-1,j}(z)) + \\ &+ zuy_k(zu) H_k(z,u) \end{aligned} \tag{2.20}$$

and by Abel's summation formula

$$\begin{aligned} \sum_{j\geq 0} u^j (y_{k,j} - y_{k-1,j}) &= -\sum_{j \geq 0} ((y_{k,j+1} - y_{k,j}) - (y_{k-1,j+1} - y_{k-1,j})) \frac{u^{j+1}-1}{u-1} \\ &= \sum_{j\geq 0} (\langle y_k, z^j\rangle z^j - \langle y_{k-1}, z^j\rangle z^j) \frac{1-u^{j+1}}{1-u} \\ &= \frac{1}{1-u}(\Delta_k(z,u) - \Delta_{k-1}(z,u)) \end{aligned}$$

with

$$\Delta_i(z,u) := y_i(z) - u y_i(zu). \tag{2.21}$$

Using this definition and observing (2.2) we further have

$$zy_k(z) + zuy_k(zu) = \frac{zy_k^2(z) - zu^2 y_k^2(zu)}{y_k(z) - uy_k(zu)} = \frac{\Delta_k(z,u) - \Delta_{k-1}(z,u)}{\Delta_k(z,u)}$$

so that (2.20) may be rewritten in the form

$$H_k(z,u)\frac{\Delta_{k-1}(z,u)}{\Delta_k(z,u)} = H_{k-1}(z,u) + \frac{1}{1-u}(\Delta_k(z,u) - \Delta_{k-1}(z,u)) \tag{2.22}$$

for all $k \geq 1$, with $H_0(z,u) = 0$.

This leads to the following explicit expression for $H_k(z,u)$:

$$H_k(z,u) = \frac{\Delta_k(z,u)}{1-u} \sum_{i=1}^{k} \frac{\Delta_i(z,u) - \Delta_{i-1}(z,u)}{\Delta_{i-1}(z,u)}. \tag{2.23}$$

In the following we denote by $\mathcal{O}_u(q_k - z)$ a formal power series in $u$, the coefficients $f_j(z)$ of which are functions in $z$ behaving like $f_j(z) = \mathcal{O}(q_k - z)$, $(z \to q_k^-)$. (The $\mathcal{O}$-constant may depend on $j$.) In this notation we have by (2.13)

$$\Delta_k(z,u) = \Delta_k(q_k,u) - a_k(q_k - z)^{1/2} + \mathcal{O}_u(q_k - z),$$
$$\Delta_i(z,u) = \Delta_i(q_k,u) + \mathcal{O}_u(q_k - z) \qquad (0 \leq i \leq k-1)$$

and

$$\frac{\Delta_k(z,u)}{\Delta_{k-1}(z,u)} = \frac{\Delta_k(q_k,u)}{\Delta_{k-1}(q_k,u)} - \frac{a_k}{\Delta_{k-1}(q_k,u)}(q_k - z)^{1/2} + \mathcal{O}_u(q_k - z).$$

Furthermore

$$H_{k-1}(z,u) = H_{k-1}(q_k,u) + \mathcal{O}_u(q_k - z)$$

which follows from the behaviour of $\Delta_i$ $(0 \leq i \leq k-1)$ and (2.23) for $H_{k-1}$. Putting everything together, (2.22) yields

$$\begin{aligned} H_k(z,u) = H_k(q_k,u) &- \frac{H_{k-1}(q_k,u)}{\Delta_{k-1}(q_k,u)} a_k (q_k - z)^{1/2} \\ &- \frac{2}{1-u} \frac{\Delta_k(q_k,u)}{\Delta_{k-1}(q_k,u)} a_k (q_k - z)^{1/2} \\ &+ \frac{1}{1-u} a_k (q_k - z)^{1/2} + \mathcal{O}_u(q_k - z). \end{aligned}$$

Observing $y_k(q_k) = 1/2q_k$ and $y_{k-1}(q_k) = 1/4q_k$ (compare (2.2)) we have

$$\begin{aligned} \Delta_{k-1}(q_k,u) &= y_{k-1}(q_k) - uy_{k-1}(q_k,u) \\ &= y_k(q_k) - y_{k-1}(q_k) + uy_k(q_k u) - uy_{k-1}(q_k u) - uy_k(q_k u) \end{aligned}$$

$$= q_k y_k^2(q_k) + q_k u^2 y_k^2(q_k u) - 2q_k y_k(q_k) u y_k(q_k u)$$
$$= q_k \Delta_k^2(q_k, u) \tag{2.24}$$

and therefore (2.19) yields with the estimation from above

$$A_k(u) = \frac{H_{k-1}(q_k, u)}{\Delta_{k-1}(q_k, u)} + \frac{1}{1-u}\left(\frac{2\Delta_k(q_k, u)}{\Delta_{k-1}(q_k, u)} - 1\right)$$
$$= \frac{H_{k-1}(q_k, u)}{q_k \Delta_k^2(q_k, u)} + \frac{1}{1-u}\left(\frac{2}{q_k \Delta_k(q_k, u)} - 1\right). \tag{2.25}$$

To give some examples we evaluate (2.25) in the cases $k=1$ and $k=2$ explicitly: The case $k=1$ corresponds to *unlabelled binary trees:* Here we have

$$H_0(z, u) = 0, \qquad \Delta_0(z, u) = 1-u, \qquad y_1(z) = (1-(1-4z)^{1/2})/2z$$
$$\Delta_1(q_1, u) = \Delta_1(\tfrac{1}{4}, u) = 2(1-u)^{1/2}$$

so that

$$A_1(u) = 4(1-u)^{-3/2} - (1-u)^{-1}. \tag{2.26}$$

The expansion yields

$$\alpha_1(j) = 2 \cdot 4^{-j}(j+1)\binom{2j+2}{j+1} - 1 \tag{2.27}$$

which is just the result of Ruskey [10] cited in the introduction.

Consider now $k=2$, which is *the case of monotone Boolean labelling:* We have to compute

$$H_1(q_2, u) = \frac{\Delta_1(q_2, u)}{1-u}\left(\frac{\Delta_1(q_2, u)}{\Delta_0(q_2, u)} - 1\right).$$

Now $\Delta_0(q_2, u) = 1-u, q_2 = \frac{3}{16}$, $y_1(q_2) = \frac{4}{3}$ and therefore

$$\Delta_1(q_2, u) = \tfrac{4}{3} - \tfrac{8}{3}(1-(1-\tfrac{3}{4}u)^{1/2}) = -\tfrac{4}{3} + \tfrac{8}{3}(1-\tfrac{3}{4}u)^{1/2}$$

which gives

$$H_1(q_2, u) = \frac{\Delta_1(q_2, u)}{(1-u)^2}(-\tfrac{4}{3} + \tfrac{8}{3}(1-\tfrac{3}{4}u)^{1/2} - (1-u)).$$

By formula (2.25)

$$A_2(u) = \frac{H_1(q_2, u)}{\Delta_1(q_2, u)} + \frac{1}{1-u}\left(\frac{2\Delta_2(q_2, u)}{\Delta_1(q_2, u)} - 1\right).$$

By (2.2) we get

$$\Delta_2(q_2, u) = \tfrac{8}{3}(-1 + 2(1-\tfrac{3}{4}u)^{1/2})^{1/2}$$

and so

$$A_2(u) = \frac{1}{(1-u)^2}(-\tfrac{4}{3} + \tfrac{8}{3}(1-\tfrac{3}{4}u)^{1/2} - (1-u))$$

$$+\frac{1}{1-u}(4(-1+2(1-\tfrac{3}{4}u)^{1/2})^{-1/2}-1)$$

or

$$A_2(u)=\frac{1}{1-u}\left(4(-1+2(1-\tfrac{3}{4}u)^{1/2})^{-1/2}+\tfrac{4}{3}\frac{1}{1-u}(-1+2(1-\tfrac{3}{4}u)^{1/2})-2\right). \tag{2.28}$$

Now for $u\to 1-$

$$-1+2(1-\tfrac{3}{4}u)^{1/2}=-1+(1+3(1-u))^{1/2}=\tfrac{3}{2}(1-u)+\mathcal{O}((1-u)^2)$$

so that we get

$$A_2(u)=4(\tfrac{2}{3})^{1/2}(1-u)^{-3/2}+\mathcal{O}((1-u)^{-1/2}) \tag{2.29}$$

and by Darboux's theorem cited above

$$\alpha_2(j)=8\left(\frac{2}{3\pi}\right)^{1/2}j^{1/2}+\mathcal{O}(j^{-1/2}) \qquad (j\to\infty). \tag{2.30}$$

Increasing values of $k$ lead to more and more complicated terms for $A_k(u)$, but it is possible to establish the asymptotic behaviour of the coefficients $\alpha_k(j)$ for $j\to\infty$ in the following way: The functions $\Delta_k(q_k,u)=y_k(q_k)-uy_k(q_k u)$ have radius of convergence 1 and the behaviour near the algebraic singularity $u=1$ is determined by the behaviour of $y_k(z)$ near $z=q_k$ (compare (2.12)):

$$\Delta_k(q_k,u)=a_k q_k^{1/2}(1-u)^{1/2}+\mathcal{O}(1-u),$$
$$\Delta_{k-1}(q_k,u)=q_k\Delta_k^2(q_k,u)=a_k^2q_k^2(1-u)+\mathcal{O}((1-u)^{3/2}) \qquad \text{(compare (2.24))}$$

and

$$H_{k-1}(q_k,u)=H_{k-1}(q_k,1)+\mathcal{O}(1-u) \qquad (u\to 1-),$$

which follows from (2.23) for $H_{k-1}$.

Putting everything together we get the following behaviour of $A_k(u)$ for $u\to 1-$ ($u=1$ must be its singularity nearest to the origin by (2.25) and the considerations above):

$$A_k(u)=2a_k^{-1}q_k^{-3/2}(1-u)^{-3/2}+\mathcal{O}((1-u)^{-1}). \tag{2.31}$$

Making again use of Darboux's theorem we get the desired asymptotic formula

$$\alpha_k(j)=4\pi^{-1/2}a_k^{-1}q_k^{-3/2}j^{1/2}+\mathcal{O}(1) \qquad (j\to\infty). \tag{2.32}$$

The constants $a_k$ can be expressed explicitly in terms of the singularities $(q_i)_{1\le i\le k}$ by formulas (6.15), (6.19) and (6.12) of Prodinger/Urbanek [9]:

$$a_k=q_k^{-1}\prod_{i=1}^{k-1}(1-2q_i)^{-1/2} \tag{2.33}$$

and therefore

$$\alpha_k(j) \sim j^{1/2} 4(\pi q_k)^{-1/2} \prod_{i=1}^{k-1} (1-2q_i)^{1/2}. \tag{2.34}$$

The asymptotic behaviour of the constants

$$C_k = 4(\pi q_k)^{-1/2} \prod_{i=1}^{k-1} (1-2q_i)^{1/2} \tag{2.35}$$

follows now by [9, (6.15) and (6.20)] to be

$$C_k \sim C k^{-1/2} \qquad (k \to \infty). \tag{2.36}$$

(Table 2 at the end of Section 3 shows the values of the constants $C_k$ for small values of $k$ under the notion $C_k(2)$.)

We summarize the main results of this section in the following theorem.

**Theorem 2.1.** *The average height of the $j+1$-st leaf of a binary tree with exactly $n$ internal nodes labelled monotonically with elements of $\{1<2<\cdots<k\}$ converges for fixed $j$ and $n \to \infty$ to a finite limit $\alpha_k(j)$ which has the following asymptotic behaviour:*

$$\alpha_k(j) \sim j^{1/2} 4(\pi q_k)^{-1/2} \prod_{i=1}^{k-1} (1-2q_i)^{1/2} \qquad (j \to \infty)$$

*where the numbers $q_i$ fulfill the recursion*

$$q_1 = \tfrac{1}{4}, \qquad q_{i+1} = q_i(1-q_i) \qquad (i \geq 1)$$

*and the constants*

$$4(\pi q_k)^{-1/2} \prod_{i=1}^{k-1} (1-2q_i)^{1/2} \sim C k^{-1/2} \qquad (k \to \infty).$$

So the order of increasement of $\alpha_k(j)$ with $j \to \infty$ is $j^{1/2}$ for all $k$, but the increasement gets smaller in this order if the set of labels is enlarged.

Of course (2.36) does not answer the question on the order of increasement of $\alpha_k(j)$ for $k \to \infty$ with a fixed value of $j$. We investigate this problem for $j=0$, that means, we compute the *average height of the leftmost leaf for 'large n':* By (2.25)

$$A_k(0) = H_{k-1}(q_k, 0) \frac{1}{q_k y_k^2(q_k)} + \left( \frac{2}{q_k y_k(q_k)} - 1 \right). \tag{2.37}$$

Now (2.23) shows that

$$H_{k-1}(q_k, 0) = y_{k-1}(q_k) \sum_{i=1}^{k-1} \frac{y_i(q_k) - y_{i-1}(q_k)}{y_{i-1}(q_k)}. \tag{2.38}$$

The values of $y_i(q_k)$ are contained implicitly in the considerations of [9, §6], but it is not difficult to compute them directly in the following way: We first extend recursion (2.11) to

$$q_0 = \tfrac{1}{2}, \qquad q_{i+1} = q_i(1-q_i) \qquad (i \geq 0) \tag{2.39}$$

and define numbers $b_i$ $(0 \leq i \leq k)$ by

$$b_i := q_k y_i(q_k).$$

Then $b_i(1-b_i) = b_{i-1}$ $(1 \leq i \leq k)$, $b_k = \frac{1}{2} = q_0$ (compare (2.2)). So we have $b_i = q_{k-i}$ $(0 \leq i \leq k)$ or

$$y_i(q_k) = \frac{q_{k-i}}{q_k} \qquad (0 \leq i \leq k). \tag{2.40}$$

Relation (2.38) reads now

$$H_{k-1}(q_k, 0) = \frac{1}{4q_k} \sum_{i=1}^{k-1} \frac{q_i}{1-q_i}$$

and by (2.37) we derive

$$A_k(0) = 3 + \sum_{i=1}^{k-1} \frac{q_i}{1-q_i} \qquad (k \geq 1). \tag{2.41}$$

The asymptotic behaviour of the sequence $(q_k)$ is established in de Bruijn [1, p. 154ff.] to be

$$q_k = \frac{1}{k} + \mathcal{O}\left(\frac{\log k}{k^2}\right) \qquad (k \to \infty) \tag{2.42}$$

and thereby

$$A_k(0) = \log k + \mathcal{O}(1) \qquad (k \to \infty). \tag{2.43}$$

**Theorem 2.2.** *The average height of the leftmost leaf of a binary tree with n nodes labelled monotonically of* $\{1<2<\cdots<k\}$ *fulfills 'for large n'*

$$\alpha_k(0) = 3 + \sum_{i=1}^{k-1} \frac{q_i}{1-q_i} \qquad (k \geq 1)$$

*with* $q_i$ *as in Theorem* 2.1. *For* $k \to \infty$

$$\alpha_k(0) \sim \log k.$$

The following Table 1 shows the first values of $\alpha_k(0)$.

Table 1

| $k$ | $\alpha_k(0)$ | $k$ | $\alpha_k(0)$ |
|---|---|---|---|
| 1 | 3 | 6 | 4.018819 |
| 2 | 3.333333 | 7 | 4.129698 |
| 3 | 3.564103 | 8 | 4.228418 |
| 4 | 3.743826 | 9 | 4.317477 |
| 5 | 3.892110 | 10 | 4.398663 |

By similar but somewhat more complicated considerations we get

**Theorem 2.3.** $\alpha_k(j) \sim C(j) \log k$ (*fixed* $j$, $k \to \infty$).

## 3. $T$-ary trees

The methods of Section 2 allow a treatment of $t$-ary trees for general $t \geq 2$. For the sake of brevity we give only short comments and refer to the more detailed proofs of the corresponding results in Section 2.

The generating functions of monotonically labelled $t$-ary trees fulfill (compare [9, 7.2])

$$y_k = y_{k-1} + z y_k^t \qquad (k \geq 1), \qquad y_0 = 1. \tag{3.1}$$

For the following observe that a $t$-ary tree with $n$ (internal) nodes has exactly $1 + (t-1)n$ leaves.

Let $H_{k,j}(z)$ again be the generating function of the sum of heights of the $j+1$-st leaf over all $t$-ary trees with $n$ (int.) nodes ($1 + (t-1)n \geq j+1$, which ensures that there are at least $j+1$ leaves). We have

$$H_{k,j} = \sum_{h \geq 0} (y_{k,j} - v_{k,j}^{[h]}) \tag{3.2}$$

where $y_{k,j}$ and $v_{k,j}^{[h]}$ have a meaning as in Section 2, especially

$$y_{k,j} = \sum_{n \geq \lfloor j/(t-1) \rfloor} \langle y_k, z^n \rangle z^n. \tag{3.3}$$

The recursions corresponding to (2.8) and (2.9) read now ($r+1$ is the number of the subtree of the root which contains the $j+1$-st leaf):

$$y_{k,j} = y_{k-1,j} + z \sum_{r=0}^{t-1} y_k^{t-1-r} \sum_{i=0}^{j-1} [y_k^r]_i y_{k,j-1-i} \tag{3.4}$$

and

$$v_{k,j}^{[h+1]} = v_{k-1,j}^{[h+1]} + z \sum_{r=0}^{t-1} y_k^{t-1-r} \sum_{i=0}^{j-1} [y_k^r]_i v_{k,j-1-i}^{[h]} \tag{3.5}$$

where

$$[y_k^r]_i = \begin{cases} z^n \langle y_k^r, z^n \rangle & \text{if } i+1 = r + (t-1)n, \\ 0 & \text{otherwise} \end{cases}$$

is the generating function of $r$-tuples of $t$-ary trees with together exactly $i+1$ leaves.

Subtracting (3.5) from (3.4) and summing up we get

$$H_{k,j}(1 - z y_k^{t-1}) = H_{k-1,j} + (y_{k,j} - y_{k-1,j}) + z \sum_{r=1}^{t-1} y_k^{t-1-r} \sum_{i=0}^{j-1} [y_k^r]_i H_{k,j-1-i}. \tag{3.6}$$

By the results of [9, §7] on the singularities of the functions $y_k$ for $t$-ary trees we have again the behaviour

$$\begin{aligned} y_{k,j}(z) &= y_{k,j}(q_k) - a_k(q_k - z)^{1/2} + \mathcal{O}(q_k - z), \\ H_{k,j}(z) &= H_{k,j}(q_k) - a_{k,j}(q_k - z)^{1/2} + \mathcal{O}(q_k - z) \end{aligned} \tag{3.7}$$

where $a_k$, $a_{k,j}$ and the singularities $q_k$ clearly depend on $t$! So the desired average heights for $n \to \infty$ are again given by

$$\alpha_k(t, j) = a_{k,j}/a_k. \tag{3.8}$$

Let

$$A_k(u) = \sum_{j \geq 0} \alpha_k(t, j) u^j \tag{3.9}$$

be the corresponding generating function. Then we have, as in (2.19),

$$a_k A_k(u) = \lim_{z \to q_k} \frac{H_k(q_k, u) - H_k(z, u)}{(q_k - z)^{1/2}} \tag{3.10}$$

with

$$H_k(z, u) = \sum_{j \geq 0} H_{k,j}(z) u^j. \tag{3.11}$$

Recursion (3.6) and Abel's summation formula (compare the proof of (2.22)) yield now

$$H_k(z, u) \frac{\Delta_{k-1}(z, u)}{\Delta_k(z, u)} = H_{k-1}(z, u) + \frac{1}{1-u}(\Delta_k(z, u) - \Delta_{k-1}(z, u)) \tag{3.12}$$

with

$$\Delta_i(z, u) := y_i(z) - u y_i(u^{t-1} z). \tag{3.13}$$

By (3.10) and (3.7) and the last definition we have again

$$A_k(u) = \frac{H_{k-1}(q_k, u)}{\Delta_{k-1}(q_k, u)} + \frac{1}{1-u}\left(\frac{2\Delta_k(q_k, u)}{\Delta_{k-1}(q_k, u)} - 1\right). \tag{3.14}$$

In order to study the asymptotic behaviour of $\alpha_k(t, j)$, $j \to \infty$, we have to investigate the behaviour of $\Delta_k(q_k, u)$ and $\Delta_{k-1}(q_k, u)$, $u \to 1-$:

$$\Delta_k(q_k, u) = y_k(q_k) - u y_k(u^{t-1} q_k) = a_k q_k^{1/2} (t-1)^{1/2} (1-u)^{1/2} + \mathcal{O}(1-u)$$

and

$$\begin{aligned} \Delta_{k-1}(q_k, u) &= y_{k-1}(q_k) - y_{k-1}(u^{t-1} q_k) + (1-u) y_{k-1}(u^{t-1} q_k) \\ &= ((t-1) q_k y'_{k-1}(q_k) + y_{k-1}(q_k))(1-u) + \mathcal{O}((1-u)^2) \end{aligned}$$

as well as

$$H_{k-1}(q_k, u) = H_{k-1}(q_k, 1) + \mathcal{O}(1-u).$$

By relation (3.14) we get

$$A_k(u) = \frac{1}{(1-u)^{3/2}} \frac{2 a_k q_k^{1/2} (t-1)^{1/2}}{(t-1) q_k y'_{k-1}(q_k) + y_{k-1}(q_k)} + \mathcal{O}((1-u)^{-1}) \tag{3.15}$$

and by the theorem of Darboux

$$\alpha_k(t,j) \sim j^{1/2}\left(\frac{t-1}{\pi}\right)^{1/2} \frac{a_k}{(t-1)q_k^{1/2}y'_{k-1}(q_k)+q_k^{-3/2}y_{k-1}(q_k)} \tag{3.16}$$

where all general constants clearly depend on $t$.

The asymptotic behaviour of the singularities $q_k$ is given by [9, (7.10)]:

$$\frac{1}{q_k} = (t-1)k + \tfrac{1}{2}t\log k + C + \mathcal{O}\left(\frac{\log k}{k}\right) \qquad (k\to\infty) \tag{3.17}$$

with a constant $C$ depending on $t$. It is possible to express all general constants of formula (3.16) by the numbers $q_i$ in the following way: By [4]

$$\tfrac{1}{2}a_k^2 = \lim_{z\to q_k^-} y'_k(z)(y_k(q_k)-y_k(z)). \tag{3.18}$$

If we remember the definition of $a_k$ and take notice of

$$y'_k(z)(1-zty_k(z)^{t-1}) = y'_{k-1}(z) + y_k(z)^t \tag{3.19}$$

(compare (3.1)), we get

$$\tfrac{1}{2}a_k^2 = \frac{y'_{k-1}(q_k)+y_k(q_k)^t}{q_k t(t-1)\, y_k(q_k)^{t-2}}. \tag{3.20}$$

By a similar proof as of (2.38) it follows from [9, (7.9)] that

$$y_i(q_k) = \left(\frac{q_{k-i}}{q_k}\right)^{1/(t-1)} \qquad (0\le i\le k), \quad \text{with } q_0 = \frac{1}{t}, \tag{3.21}$$

especially

$$y_k(q_k) = (tq_k)^{-1/(t-1)}. \tag{3.22}$$

It remains to compute $y'_{k-1}(q_k)$: Let $c_i := y'_i(q_k)$ $(0\le i\le k-1)$. Relation (3.19) with $k$ substituted by $i$ and $z=q_k$ yields

$$c_i(1-q_k t y_i(q_k)^{t-1}) = c_{i-1} + y_i(q_k)^t$$

and because of (3.21) this means

$$c_i(1-tq_{k-i}) = c_{i-1} + \left(\frac{q_{k-i}}{q_k}\right)^{t/(t-1)}$$

so that finally

$$y'_{k-1}(q_k) = \sum_{r=1}^{k-1}\left(\frac{q_r}{q_k}\right)^{t/(t-1)} \prod_{i=1}^{r}\frac{1}{1-tq_i}. \tag{3.23}$$

By the asymptotic formula (3.17) (i.e. the methods of [1, p. 154ff.]) and Satz 10 of Knopp [5, p. 231] it is not difficult to prove that

$$y'_{k-1}(q_k) \sim C'(t)k^{t/(t-1)}$$

as well as

$$a_k \sim C''(t)k^{t/(t-1)} \qquad (k \to \infty)$$

so that we get (a closed formula for $C_k(t)$ would be very complicated)

**Theorem 3.1.** *The average height of the $j+1$-st leaf of a $t$-ary tree with $n$ (interval) nodes labelled monotonically with elements of $\{1<2<\cdots<k\}$ converges for fixed $j$ and $n \to \infty$ to $\alpha_k(t,j)$, where*

$$\alpha_k(t,j) \sim j^{1/2}C_k(t) \qquad (j \to \infty)$$

*where the constants $C_k(t)$ are expressed by relations* (3.16), (3.20), (3.21) *and* (3.23) *and behave like*

$$C_k(t) \sim C(t)k^{-1/2} \qquad (k \to \infty).$$

Even the special case $k=1$, which corresponds to the family of *unlabelled $t$-ary trees* does not seem to have been treated in literature so far: By (3.16)

$$\alpha(t,1,j) \sim j^{1/2}4\left(\frac{t-1}{\pi}\right)^{1/2} a_1 q_1^{1/2}$$

where

$$\tfrac{1}{2}a_1^2 = \frac{y_1(q_1)^2}{q_1 t(t-1)} \qquad \text{(compare (3.20))}.$$

With

$$q_1 = \frac{1}{t}\left(\frac{t-1}{t}\right)^{t-1} \qquad \text{(compare [9, §7])}$$

$$y_1(q_1) = \frac{t}{t-1} \qquad \text{(compare (3.22))}$$

and therefore

$$a_1 = 2^{1/2}t^{(t+1)/2}(t-1)^{-(t+2)/2}$$

so that we get

**Corollary 3.2.** *For unlabelled $t$-ary trees the numbers $\alpha(t,1,j)$ described in Theorem* 3.1 *fulfill*

$$\alpha_1(t,j) \sim j^{1/2}\frac{8}{(2\pi)^{1/2}}\frac{t^{1/2}}{t-1} \qquad (j \to \infty). \tag{3.24}$$

We conclude this section with Table 2 of $C_k(t)$ for small values of $t$ and $k$: for this purpose we use the following recurrence relation, which makes it possible to compute the exact values of the numbers $q_k$ (compare [9, (7.8f)])

$$q_k = \frac{r_k^{t-1}}{t}, \quad \text{where } r_0 = 1, \qquad r_{k+1} = r_k - \frac{r_k^t}{t} \quad (k \ge 0)$$

from which

$$q_{k+1} = q_k(1-q_k)^{t-1}, \quad \text{with } q_0 = 1/t. \tag{3.25}$$

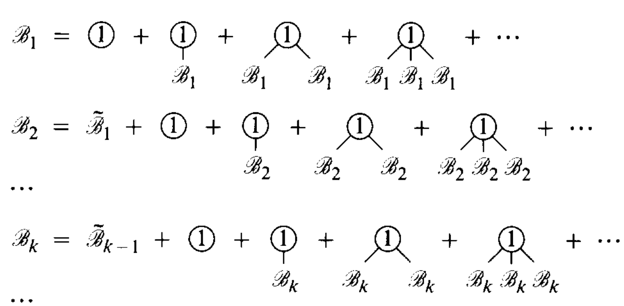
Table 2. Values of $C_k(t)$

| $k \backslash t$ | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| 1 | 4.513517 | 2.763953 | 2.127692 | 1.784124 | 1.563528 | 1.407336 |
| 2 | 3.685271 | 2.232094 | 1.710495 | 1.430731 | 1.251858 | 1.125578 |
| 3 | 3.232196 | 1.944734 | 1.486287 | 1.241377 | 1.085178 | 0.975094 |
| 4 | 2.927372 | 1.753171 | 1.337398 | 1.115901 | 0.974876 | 0.875603 |
| 5 | 2.701627 | 1.612316 | 1.228242 | 1.024059 | 0.894223 | 0.802906 |
| 6 | 2.524671 | 1.502540 | 1.143371 | 0.952740 | 0.831643 | 0.746530 |
| 7 | 2.380593 | 1.413586 | 1.074730 | 0.895120 | 0.781117 | 0.701034 |
| 8 | 2.260035 | 1.339455 | 1.017620 | 0.847222 | 0.739139 | 0.663249 |
| 9 | 2.157050 | 1.276350 | 0.969073 | 0.806535 | 0.703497 | 0.631178 |
| 10 | 2.067636 | 1.221728 | 0.927101 | 0.771383 | 0.672716 | 0.603488 |

## 4. Ordered trees

The considerations on binary and $t$-ary trees essentially bear the ideas of the treatment of this present case already if they are read with the right interpretation, and, vice versa, the considerations of this chapter will give more insight why a special style of presentation has been chosen for some of the proofs and results in Sections 2 and 3.

According to [9] and the usual definition of ordered trees in literature we have the following formal equalities for the families $\mathscr{B}_k$ of ordered trees with nodes labelled monotonically with elements of $\{1<2<\cdots<k\}$:



(4.1)

where $\tilde{\mathscr{B}}_k$ is again the family with labels taken from $\{2<\cdots<k+1\}$. The corresponding generating functions must fulfill

$$y_k = y_{k-1} + \frac{z}{1-y_k} \qquad (k \geq 1), \qquad y_0 = 0. \tag{4.2}$$

**Remark.** For the sake of analogy of the results of this section with those of 2 and 3 it is very much convenient to alter the meaning of some notations:

(a) There is *no distinction between leaves and (internal) nodes now* as it was with binary and $t$-ary trees, which means that leaves (i.e. nodes having no son) have now to be counted within the concerned number of nodes. As a consequence the height of a leaf shall now be the number of all nodes in the chain connecting the root with the leaf.

(b) The argument or index $j$ in terms concerning a special leaf will now refer to the *$j$-th leaf* (and not to the $j+1$-st as before).

Let $y_{k;n,\lambda}$ be the number of ordered trees in $\mathscr{B}_k$ with exactly $n$ nodes and $\lambda$ leaves. The generating functions corresponding to $y_k$ and $y_{k,j}$ in Sections 2 and 3 are now

$$\begin{aligned} y_{k,j}(z,u) &:= \sum_{n\geq 1,\lambda\geq j} y_{k;n,\lambda} z^n u^\lambda, \\ y_{k,j}(z) &:= y_{k,j}(z,1), \\ y_k(z,u) &:= y_{k,1}(z,u), \\ y_k(z) &= y_{k,1}(z) = y_k(z,1). \end{aligned} \tag{4.3}$$

We remark for later considerations that the functions $y_k(z,u)$ fulfill the equation

$$y_k(z,u) = y_{k-1}(z,u) + zu + \frac{zy_k(z,u)}{1-y_k(z,u)}, \tag{4.4}$$

which can be seen immediately from (4.2) and the definition. (4.4) may be transformed to the recursion

$$\begin{aligned} &y_k(z,u) = \tfrac{1}{2}(1-z(1-u)+y_{k-1}(z,u)-[(1+z(1-u)-y_{k-1}(z,u))^2-4z]^{1/2}) \\ &\qquad \text{for all } k\geq 1, \qquad y_0(z,u)=0. \end{aligned} \tag{4.5}$$

The functions $y_{k,j}(z)$ fulfill the recursion

$$\begin{aligned} y_{k,j}(z) &= y_{k-1,j}(z) + z\delta_{j,1} \\ &\quad + z\sum_{t\geq 1}\sum_{r=0}^{t-1} y_k(z)^{t-1-r}\sum_{i=0}^{j-1}\langle y_k^r(z,u), u^i\rangle y_{k,j-i}(z). \end{aligned} \tag{4.6}$$

(The first term of the right side corresponds to the trees of $\tilde{\mathscr{B}}_k$, the second term to ① and the third one to all remaining trees of $\mathscr{B}_k$, where $t$ denotes the number of subtrees of the root and the $j$-th leaf is in the $r+1$-st of these subtrees.)

Let $v_{k,j}^{[h]}(z)$ be the generating function of trees in $\mathscr{B}_k$ with at least $j$ leaves, where the height of the $j$-th leaf is smaller or equal to $h$. Just the same considerations as above lead to

$$\begin{aligned} v_{k,j}^{[h+1]}(z) &= v_{k-1,j}^{[h+1]}(z) + z\delta_{j,1} \\ &\quad + z\sum_{t\geq 1}\sum_{r=0}^{t-1} y_k(z)^{t-1-r}\sum_{i=0}^{j-1}\langle y_k^r(z,u), u^i\rangle v_{k,j-i}^{[h]}(z) \end{aligned} \tag{4.7}$$

with $v_{k,j}^{[0]}(z)=0$. Let

$$H_{k,j}(z) := \sum_{h\geq 0} (y_{k,j}(z) - v_{h,j}^{[h]}(z)) \tag{4.8}$$

be the 'sum of the heights of the $j$-th leaf'-generating function. Then (4.7) yields the recursion

$$\begin{aligned} H_{k,j}(z) &= H_{k-1,j}(z) + y_{k,j}(z) - y_{k-1,j}(z) \\ &\quad + z\sum_{t\geq 1}\sum_{r=0}^{t-1} y_k(z)^{t-1-r} \sum_{i=0}^{j-1} \langle y_k^r(z,u), u^i\rangle H_{k,j-i}(z) \end{aligned} \tag{4.9}$$

and with the definition

$$H_k(z,u) := \sum_{j\geq 1} H_{k,j}(z) u^j \tag{4.10}$$

we get

$$\begin{aligned} H_k(z,u) &= H_{k-1}(z,u) + \sum_{j\geq 1} u^j (y_{k,j}(z) - y_{k-1,j}(z)) \\ &\quad + z\sum_{t\geq 1}\sum_{r=0}^{t-1} y_k(z)^{t-1-r} y_k(z,u)^r H_k(z,u) \end{aligned}$$

or equivalently

$$\begin{aligned} H_k(z,u)\left(1 - \frac{z}{(1-y_k(z))(1-y_k(z,u))}\right) &= H_{k-1}(z,u) \\ &\quad + \sum_{j\geq 1} u^j (y_{k,j}(z) - y_{k-1,j}(z)). \end{aligned}$$

Let

$$\Delta_i(z,u) := y_i(z) - y_i(z,u). \tag{4.11}$$

With this abbreviation we have by recursions (4.2) and (4.4)

$$\begin{aligned} \frac{z}{(1-y_k(z))(1-y_k(z,u))} &= \frac{1}{\Delta_k(z,u)}\left(\frac{z}{1-y_k(z)} - \frac{z}{1-y_k(z,u)}\right) \\ &= \frac{1}{\Delta_k(z,u)}(y_k(z) - y_{k-1}(z) - y_k(z,u) \\ &\quad + y_{k-1}(z,u) - z(1-u)) \\ &= \frac{1}{\Delta_k(z,u)}(\Delta_k(z,u) - \Delta_{k-1}(z,u) - z(1-u)). \end{aligned}$$

By Abel's summation formula (compare the proof of (2.22))

$$\sum_{j\geq 1} u^j (y_{k,j}(z) - y_{k-1,j}(z)) = \frac{u}{1-u}(\Delta_k(z,u) - \Delta_{k-1}(z,u)),$$

so that we find

$$H_k(z,u)\frac{\Delta_{k-1}(z,u) + z(1-u)}{\Delta_k(z,u)} = H_{k-1} + \frac{u}{1-u}(\Delta_k(z,u) - \Delta_{k-1}(z,u)). \tag{4.12}$$

The behaviour of $y_k(z)$ near its (algebraic) singularity $q_k$ nearest to the origin has been investigated in [9, §8] and is again of the type

$$y_k(z) = y_k(q_k) - a_k(q_k - z)^{1/2} + \mathcal{O}(q_k - z) \qquad (z \to q_k^-). \tag{4.13}$$

The functions $y_k(z, u)$ behave like

$$y_k(z, u) = y_k(q_k, u) + \mathcal{O}(q_k - z) \tag{4.14}$$

with $\mathcal{O}_u$ in the sense of Section 2: Consider the singularity $(\in \mathbb{R})$ $z = p_k(u)$ of $y_k(z, u)$ nearest to the origin; it is determined by the equation

$$f(z, u) := (1 + z(1-u) - y_{k-1}(z, u))^2 - 4z = 0$$

(compare (4.5)). Now $z = q_k$ is the zero nearest to the origin of

$$f(z, 1) = (1 - y_{k-1}(z))^2 - 4z$$

and

$$f(z, u) > f(z, 1) \geq 0$$

for any fixed $0 < u < 1$ and $0 \leq z \leq q_k$, so that $p_k(u) > q_k$.

A fortiori each of the coefficient series $\langle y_k(z, u), u^j \rangle$ in $z$ must have a radius of convergence greater than $q_k$.

If we define (with $\alpha_k(j)$ as in Section 2)

$$A_k(u) := \sum_{j \geq 1} \alpha_k(j) u^j \tag{4.15}$$

we get

$$a_k A_k(u) = \lim_{z \to q_k^-} \frac{H_k(q_k, u) - H_k(z, u)}{(q_k - z)^{1/2}}. \tag{4.16}$$

Now

$$\Delta_k(z, u) = \Delta_k(q_k, u) - a_k(q_k - z)^{1/2} + \mathcal{O}_u(q_k - z)$$

$$\Delta_{k-1}(z, u) = \Delta_{k-1}(q_k, u) + \mathcal{O}_u(q_k - z)$$

so that

$$\frac{\Delta_k(z, u)}{\Delta_{k-1}(z, u) + z(1-u)} = \frac{\Delta_k(q_k, u)}{\Delta_{k-1}(q_k, u) + q_k(1-u)} - \frac{a_k}{\Delta_{k-1}(q_k, u) + q_k(1-u)} (q_k - z)^{1/2} + \mathcal{O}_u(q_k - z).$$

Relation (4.12) leads now to the estimation

$$\begin{aligned} H_k(z, u) = H_k(q_k, u) &- H_{k-1}(q_k, u) \frac{a_k}{\Delta_{k-1}(q_k, u) + q_k(1-u)} (q_k - z)^{1/2} \\ &- \frac{u}{1-u} a_k \frac{\Delta_k(q_k, u) - \Delta_{k-1}(q_k, u)}{\Delta_{k-1}(q_k, u) + q_k(1-u)} (q_k - z)^{1/2} \\ &- \frac{u}{1-u} a_k \frac{\Delta_k(q_k, u)}{\Delta_{k-1}(q_k, u) + q_k(1-u)} (q_k - z)^{1/2} \\ &+ \mathcal{O}_u(q_k - z) \end{aligned}$$

and finitely by (4.16)

$$A_k(u)=\frac{H_{k-1}(q_k,u)}{\Delta_{k-1}(q_k,u)+q_k(1-u)}+\frac{u}{1-u}\,\frac{2\Delta_k(q_k,u)-\Delta_{k-1}(q_k,u)}{\Delta_{k-1}(q_k,u)+q_k(1-u)}. \tag{4.17}$$

As in Sections 2 and 3 we consider the example $k=1$ in detail, which corresponds to *unlabelled ordered trees*. Here we have by (4.17)

$$A_1(u)=\frac{u}{1-u}\,\frac{2\Delta_1(q_1,u)}{q_1(1-u)}=\frac{8u}{(1-u)}\Delta_1(q_1,u).$$

Now $q_1=\frac{1}{4}$, $y_1(q_1)=\frac{1}{2}$ (compare (4.2)) and by (4.5)

$$\Delta_1(q_1,u)=y_1(q_1)-y_1(q_1,u)=\frac{1-u}{8}+\frac{(1-u)^{1/2}}{2\sqrt{2}}\left(1+\frac{1-u}{8}\right)^{1/2}$$

and therefore

$$\begin{aligned}A_1(u)&=\frac{u}{1-u}+\frac{2\sqrt{2}u}{(1-u)^{3/2}}\left(1+\frac{1-u}{8}\right)^{1/2}\\ &=\frac{2\sqrt{2}}{(1-u)^{3/2}}+\frac{1}{1-u}+\mathcal{O}((1-u)^{-1/2})\end{aligned} \tag{4.18}$$

and by Darboux's theorem

$$\alpha_1(j)=8(2\pi)^{-1/2}j^{1/2}+1+\mathcal{O}(j^{-1/2})\qquad (j\to\infty) \tag{4.19}$$

which is just the result of Kemp [5] on the average height of the 'MAX-turns' of a stack during postorder-traversing of a binary tree.

Next we discuss the situation for general $k$ in short: By (4.5)

$$\begin{aligned}\Delta_k(q_k,u)&=y_k(q_k)-y_k(q_k,u)\\ &=y_k(q_k)-\tfrac{1}{2}+\tfrac{1}{2}q_k(1-u)-\tfrac{1}{2}y_{k-1}(q_k,u)\\ &\quad+\tfrac{1}{2}[(1+q_k(1-u)-y_{k-1}(q_k,u))^2-4q_k]^{1/2}.\end{aligned}$$

It is an immediate consequence of (4.2) resp. the definition of $q_k$ (compare also [9, (8.6)]) that

$$y_k(q_k)=\tfrac{1}{2}(1+y_{k-1}(q_k))\quad\text{and}\quad q_k=(1-y_k(q_k))^2.$$

with these identities and the definition

$$d_{k-1}:=\left.\frac{\partial y_{k-1}}{\partial u}\right|_{u=1,z=q_k} \tag{4.20}$$

we get

$$\begin{aligned}\Delta_k(q_k,u)&=\mathcal{O}(1-u)+\tfrac{1}{2}[(1-y_{k-1}(q_k)+q_k(1-u)+d_{k-1}(1-u))^2\\ &\quad-4q_k+\mathcal{O}((1-u)^2)]^{1/2}\\ &=(q_k^{3/2}+q_k^{1/2}d_{k-1})^{1/2}(1-u)^{1/2}+\mathcal{O}(1-u)\qquad(u\to 1-).\end{aligned} \tag{4.21}$$

Furthermore

$$\Delta_{k-1}(q_k, u) = d_{k-1}(1-u) + \mathcal{O}((1-u)^2)$$

so that by

$$A_k(u) = \frac{u}{1-u} \frac{2\Delta_k(q_k, u)}{\Delta_{k-1}(q_k, u) + q_k(1-u)} + \mathcal{O}((1-u)^{-1})$$

$$A_k(u) = 2q_k^{1/4}(q_k + d_{k-1})^{-1/2}(1-u)^{-3/2} + \mathcal{O}((1-u)^{-1}) \tag{4.22}$$

and

$$\alpha_k(j) \sim j^{1/2} 4(\pi(q_k^{1/2} + q_k^{-1/2} d_{k-1}))^{-1/2} + \mathcal{O}(1) \qquad (j \to \infty). \tag{4.23}$$

The constants $d_i$ defined in (4.20) may be determined as follows: By (4.4)

$$d_i\left(1 - \frac{q_k}{(1-y_i(q_k))^2}\right) = q_k + d_{i-1}$$

and by [9, (8.6)] this is equivalent with

$$d_i(1 - q_{k-i}) = q_k + d_{i-1} \tag{4.24}$$

so that

$$d_{k-1} = q_k \sum_{r=1}^{k-1} \prod_{i=1}^{r} \frac{1}{1-q_i}. \tag{4.25}$$

In [9, §8] it is shown that the singularities $q_i$ fulfill the recursion

$$q_{i+1}^{-1} = q_i^{-1} + q_i + 2, \qquad q_1 = \tfrac{1}{4} \qquad (\text{or } q_{i+1} = q_i/(1+q_i)^2) \tag{4.26}$$

and behave like

$$q_k = \frac{1}{2k} + \mathcal{O}\left(\frac{\log k}{k^2}\right) \qquad (k \to \infty). \tag{4.27}$$

Using again Satz 10 of Knopp [6, p. 231] we find

$$\prod_{i=1}^{r} (1-q_i) \sim C' r^{-1/2} \qquad (r \to \infty)$$

and therefore

$$\sum_{r=1}^{k-1} \prod_{i=1}^{r} \frac{1}{1-q_i} \sim C'' k^{3/2} \qquad (k \to \infty)$$

which yields

$$\begin{aligned} &\alpha_k(j) \sim C_k j^{1/2} \qquad (j \to \infty) \\ &\quad \text{with } C_k \sim C k^{-1/2} \qquad (k \to \infty). \end{aligned} \tag{4.28}$$

**Theorem 4.1.** *The average height of the $j$-th leaf of an ordered tree with $n$ nodes labelled monotonically with labels of $\{1 < 2 < \cdots < k\}$ converges for fixed $j$ and*

*$n \to \infty$ to a finite limit $\alpha_k(j)$, which has the following asymptotic behaviour*

$$\alpha_k(j) \sim j^{1/2} C_k = j^{1/2} 4\left(\pi q_k^{1/2}\left(1 + \sum_{r=1}^{k-1} \prod_{i=1}^{r} \frac{1}{1-q_i}\right)\right)^{-1/2} \qquad (j \to \infty)$$

*where the numbers $q_i$ fulfill the recursion*

$$q_{i+1}^{-1} = q_i^{-1} + q_i + 2, \qquad q_1 = \tfrac{1}{4},$$

*and the constants*

$$C_k \sim C k^{-1/2} \qquad (k \to \infty).$$

The following Table 3 shows the values of $C_k$ for $1 \le k \le 10$:

Table 3

| $k$ | $C_k$ | $k$ | $C_k$ |
|---|---|---|---|
| 1 | 3.191538 | 6 | 1.406303 |
| 2 | 2.335967 | 7 | 1.307827 |
| 3 | 1.940912 | 8 | 1.227788 |
| 4 | 1.699422 | 9 | 1.161033 |
| 5 | 1.531721 | 10 | 1.104231 |

## References

[1] N.G. de Bruijn, Asymptotic Methods in Analysis (North-Holland, Amsterdam, 1958).
[2] L. Comtet, Advanced Combinatorics (D. Reidel, Dordrecht, 1974).
[3] Ph. Flajolet, Analyse d'algorithmes de manipulation d'arbres et de fichiers, Thèse Paris-Sud-Orsay (1979).
[4] F. Harary and E.M. Palmer, Graphical Enumeration (Academic Press, New York-London, 1973).
[5] R. Kemp, On the average oscillation of a stack, Combinatorica 2 (1982) 157-176.
[6] K. Knopp, Theorie und Anwendungen der unendlichen Reihen, 3. Aufl. (Springer, Berlin, 1931).
[7] D.E. Knuth, The Art of Computer Programming, Fundamental Algorithms, Vol. 1 (Addison-Wesley, New York, 1968).
[8] A. Meir and J.W. Moon, On the altitude of nodes in random trees, Canad. J. Math. 30 (1978) 997-1015.
[9] H. Prodinger and F.J. Urbanek, On monotone functions of tree structures, Discrete Appl. Math. 5 (1983) 223-239.
[10] F. Ruskey, On the average shape of binary trees, SIAM J. Algebraic Discrete Methods 1 (1980) 43-50.